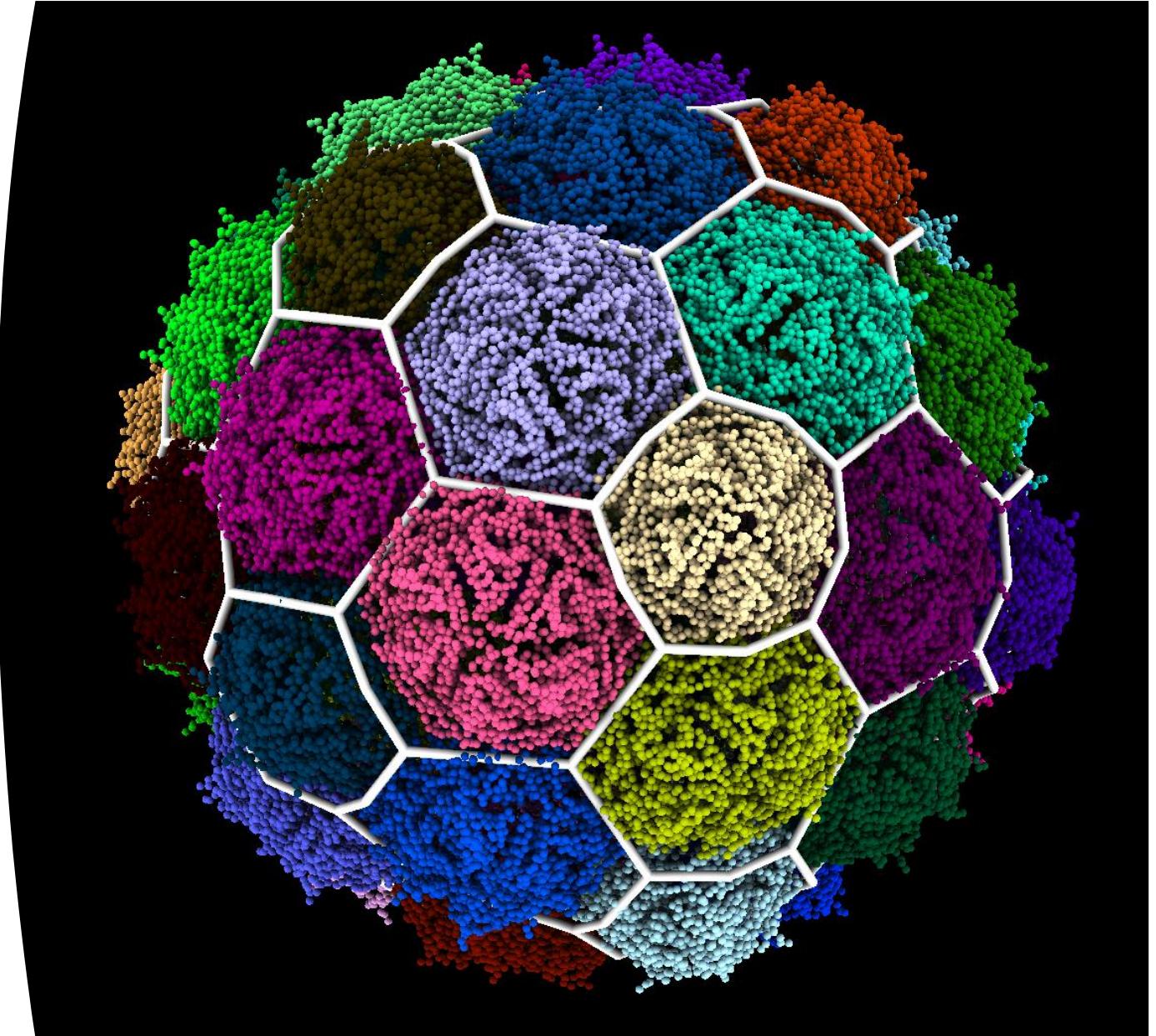
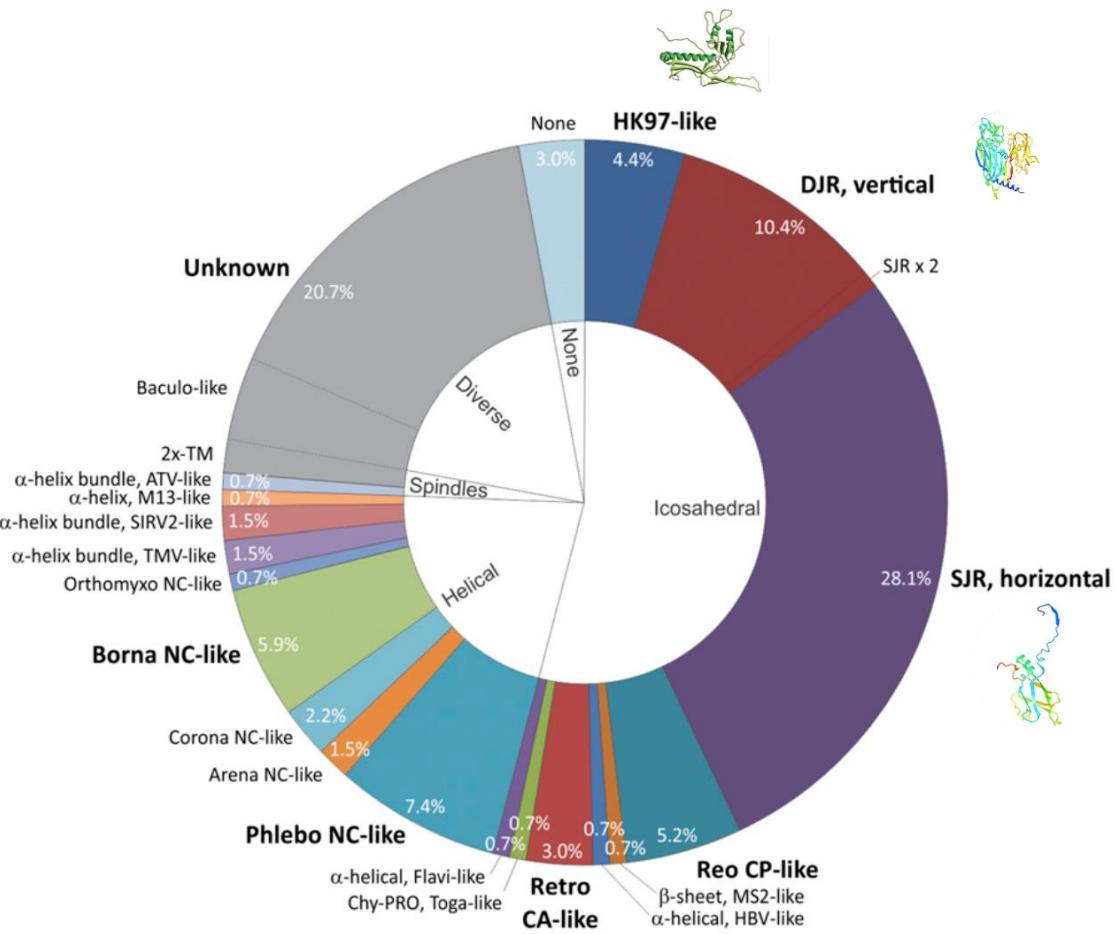


Structural Analysis & Classification Of HK97-like Viral Capsids

Colin Brown
Luque Lab



Icosahedral capsids dominate the virosphere



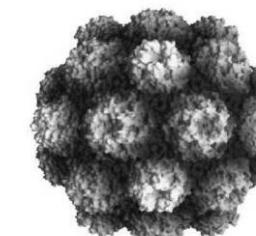
Krupovic and Koonin
PNAS 2017

Viral Capsids

- Contain the viral genome
- Built from repeated units

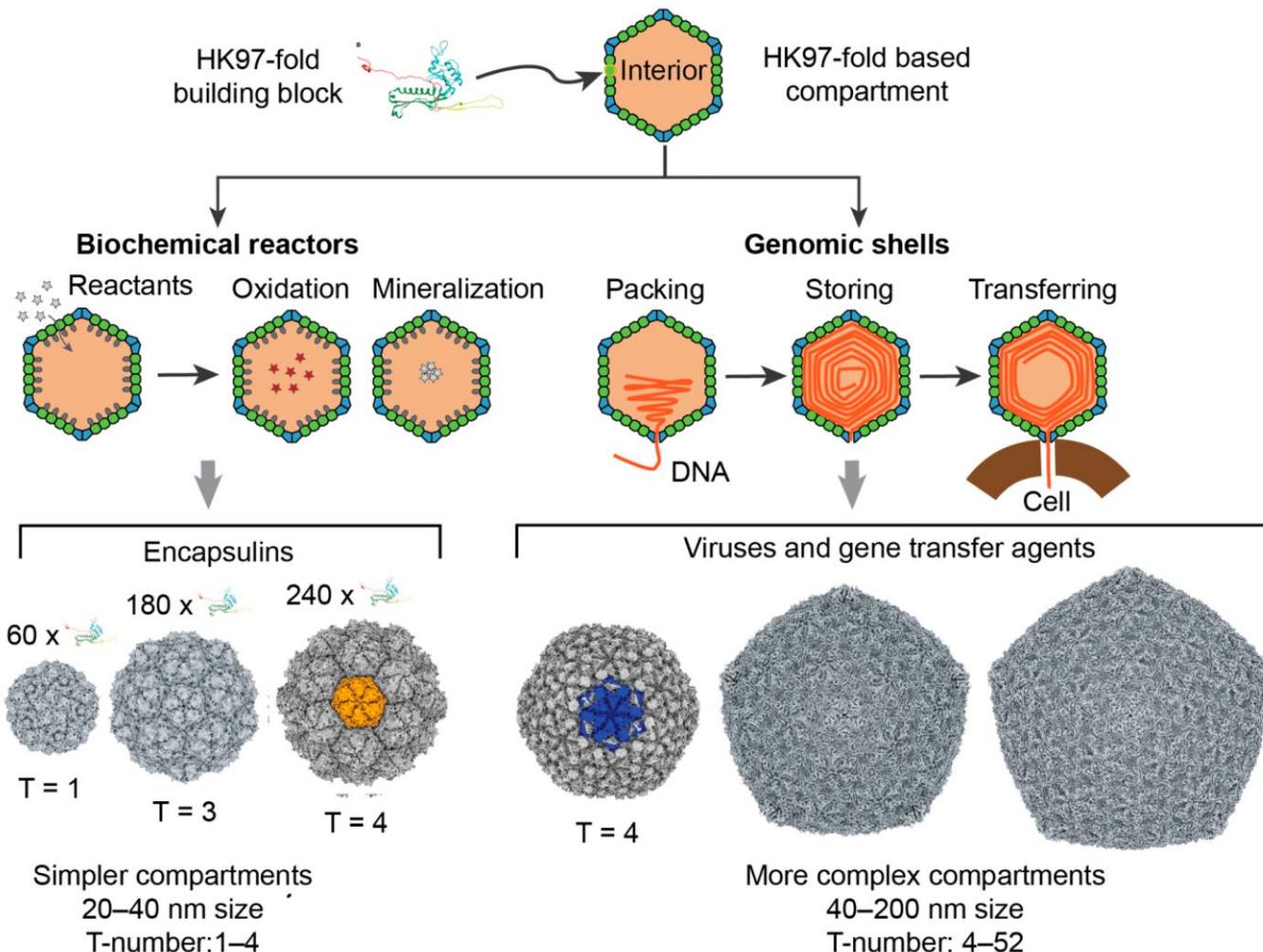
Icosahedral Capsids

- Highly symmetric
- >90% of viruses

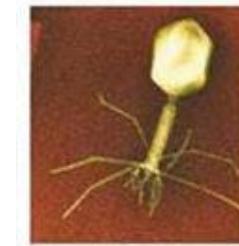


ICTVdB

HK97-Like Folds



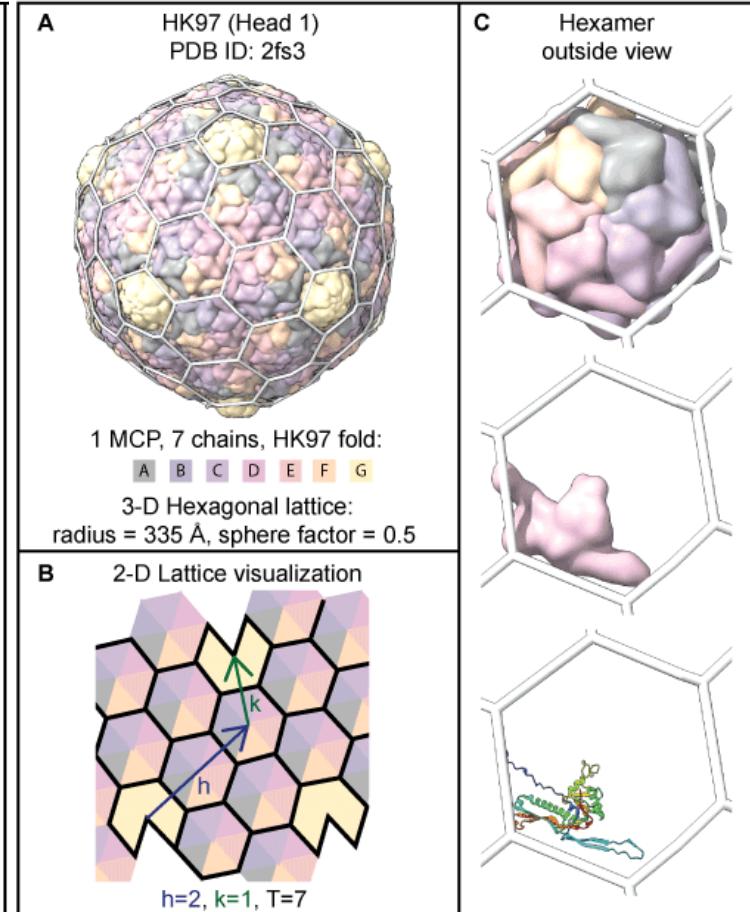
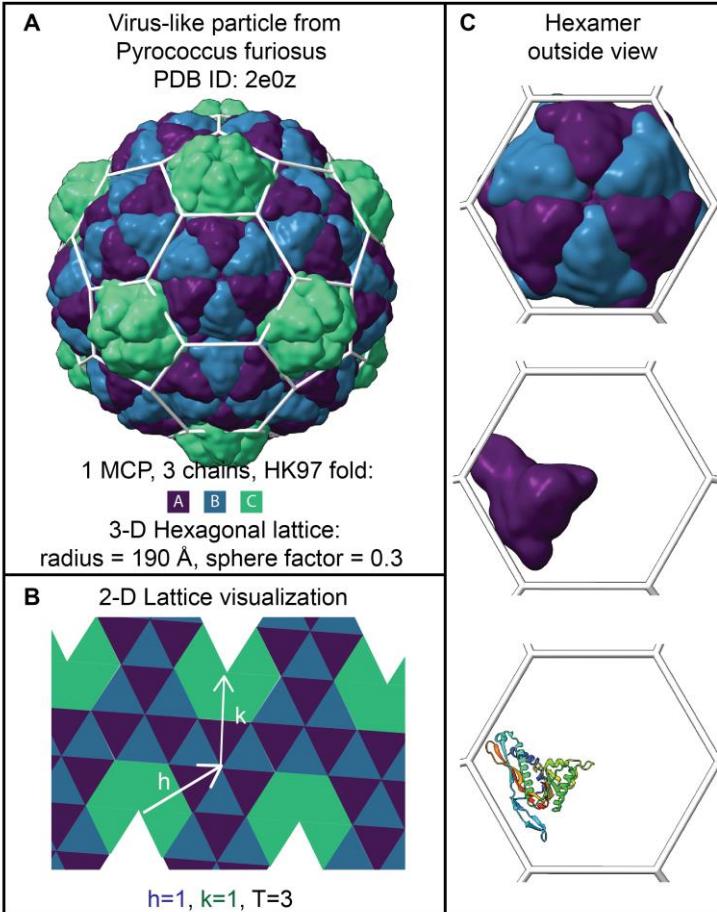
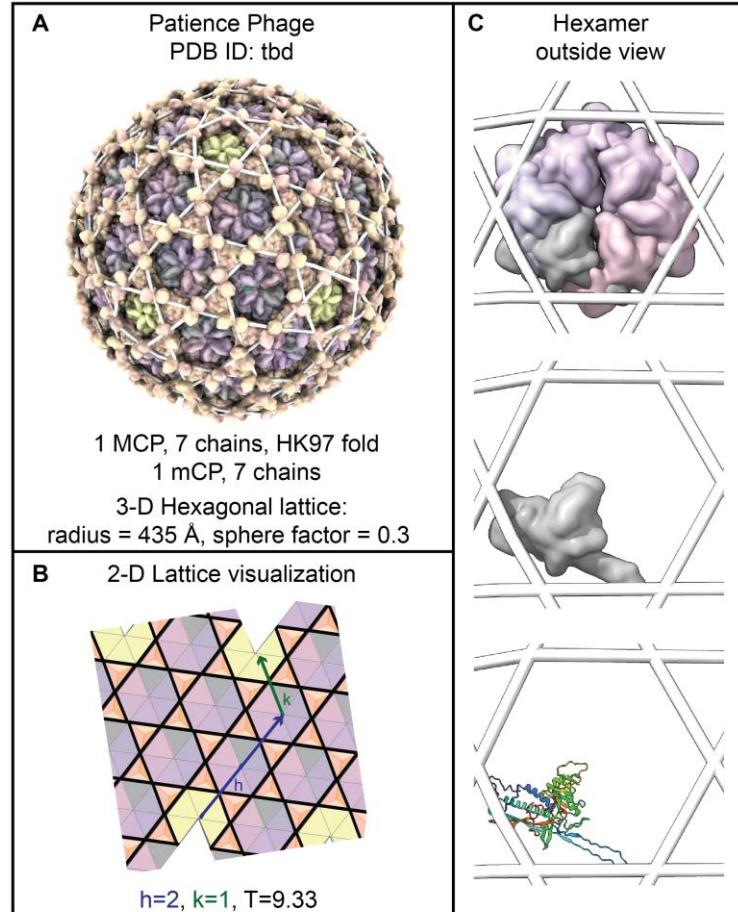
HK97-fold



Tailed bacteriophages

50%-90% of viral particles in
the environment

Classifying Icosahedral Capsid Geometry

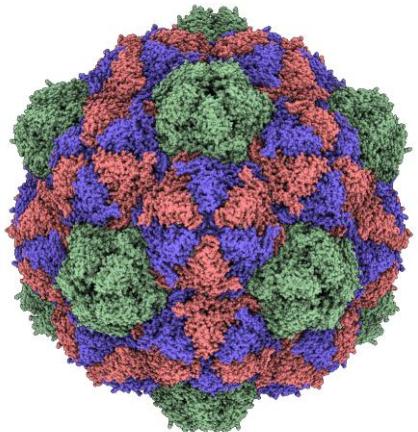


Lee, Brown, Bartels
In Preparation

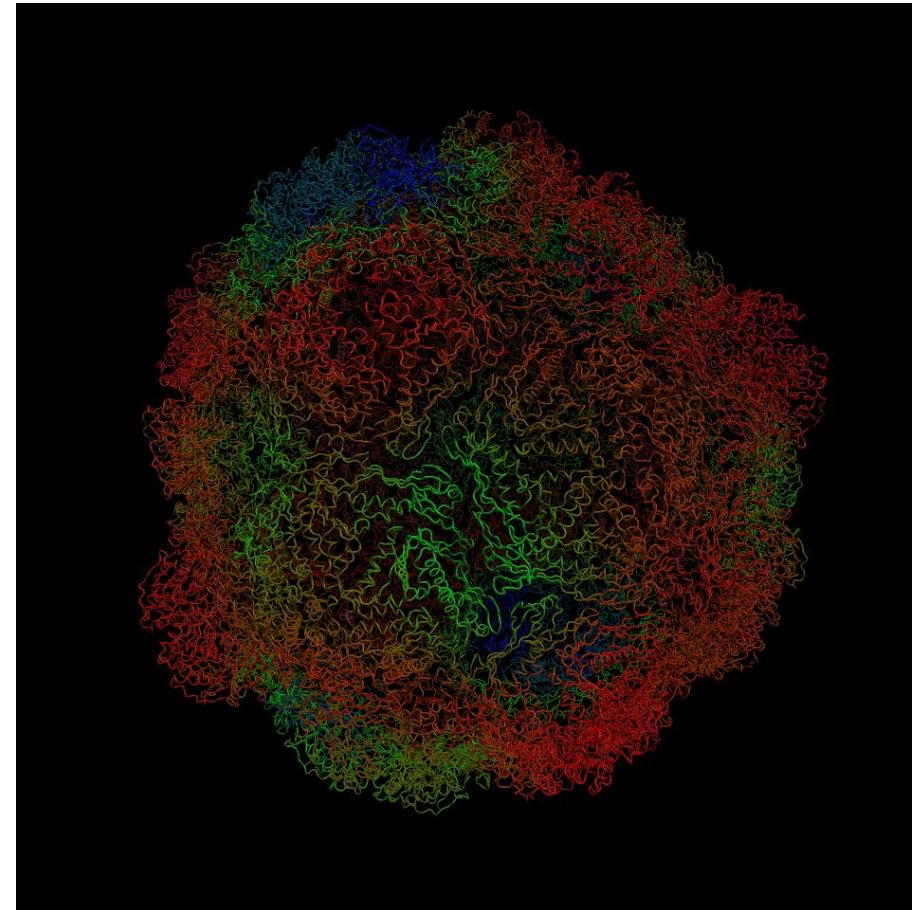
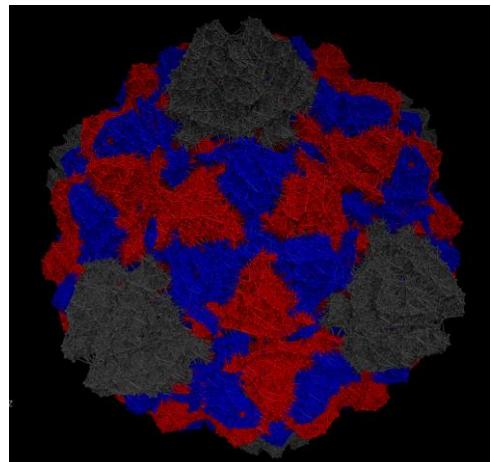
Step 1: Physical Model

Normal Mode Analysis

Atomic model

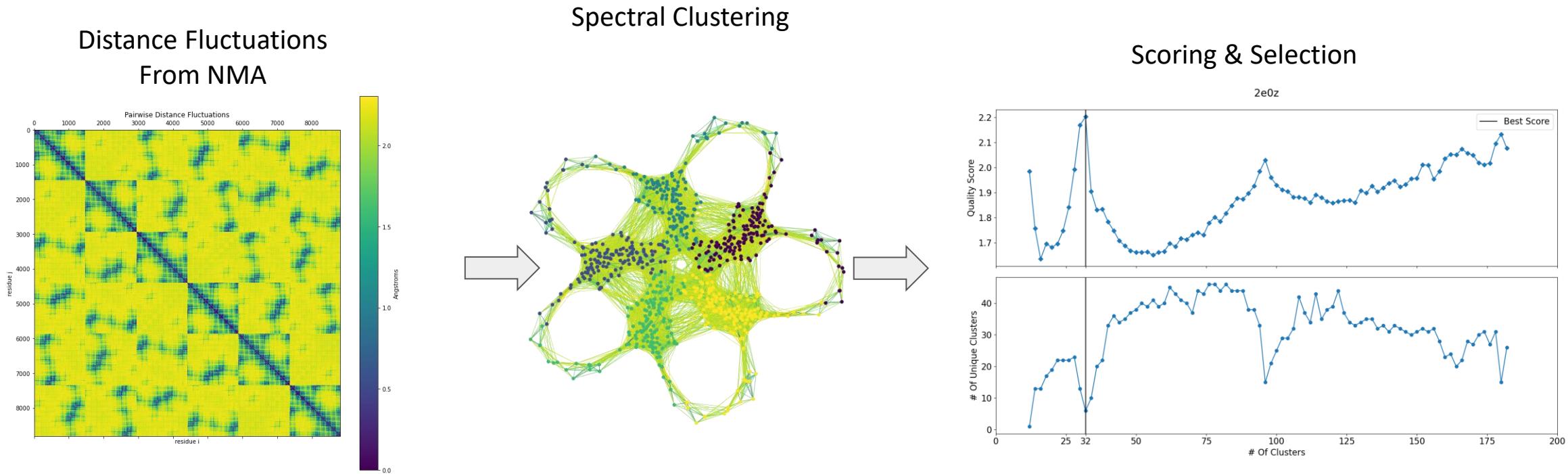


Elastic Network Model



Brown, Luque
In Preparation

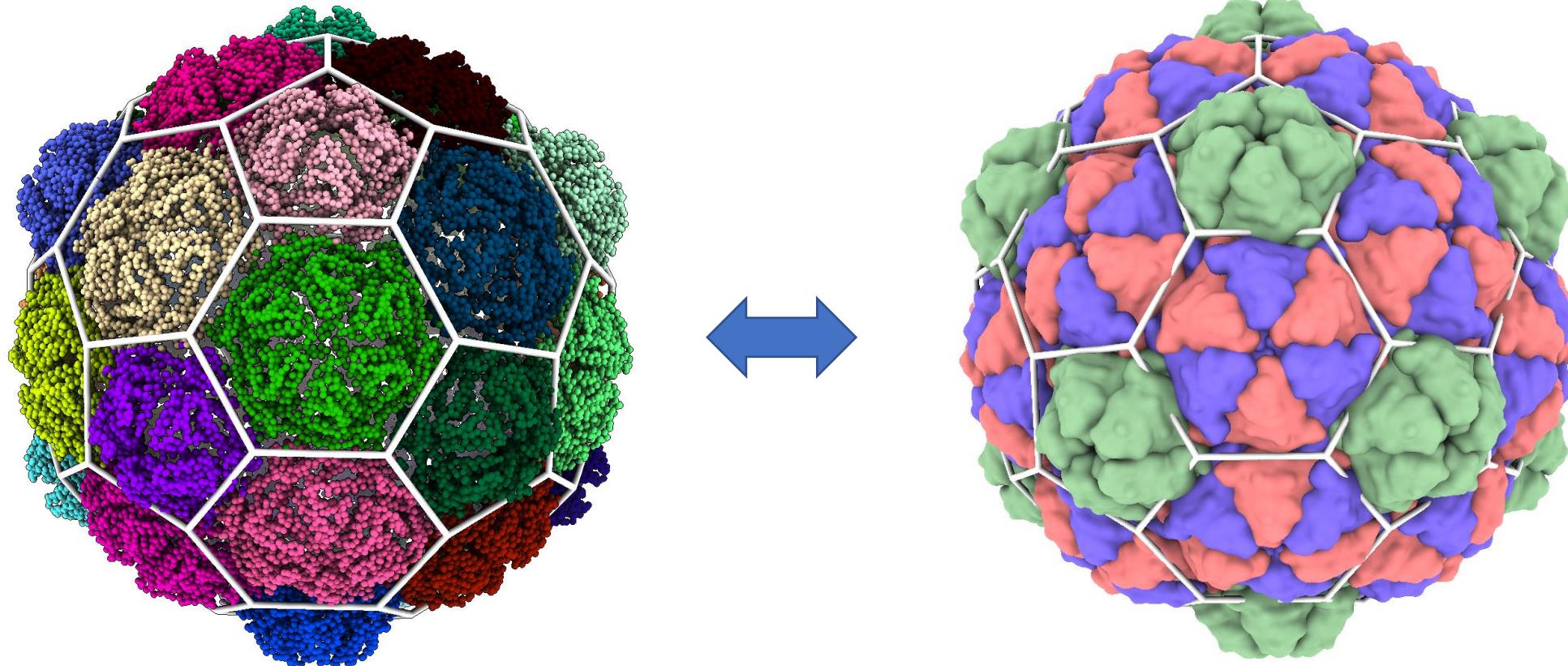
Step 2: Quasi-Rigid Subdivision



$$f_{ij}^2 = \text{Var}(d_{ij}^2) = \langle d_{ij}^2 \rangle - \langle d_{ij} \rangle^2$$

Brown, Luque
In Preparation

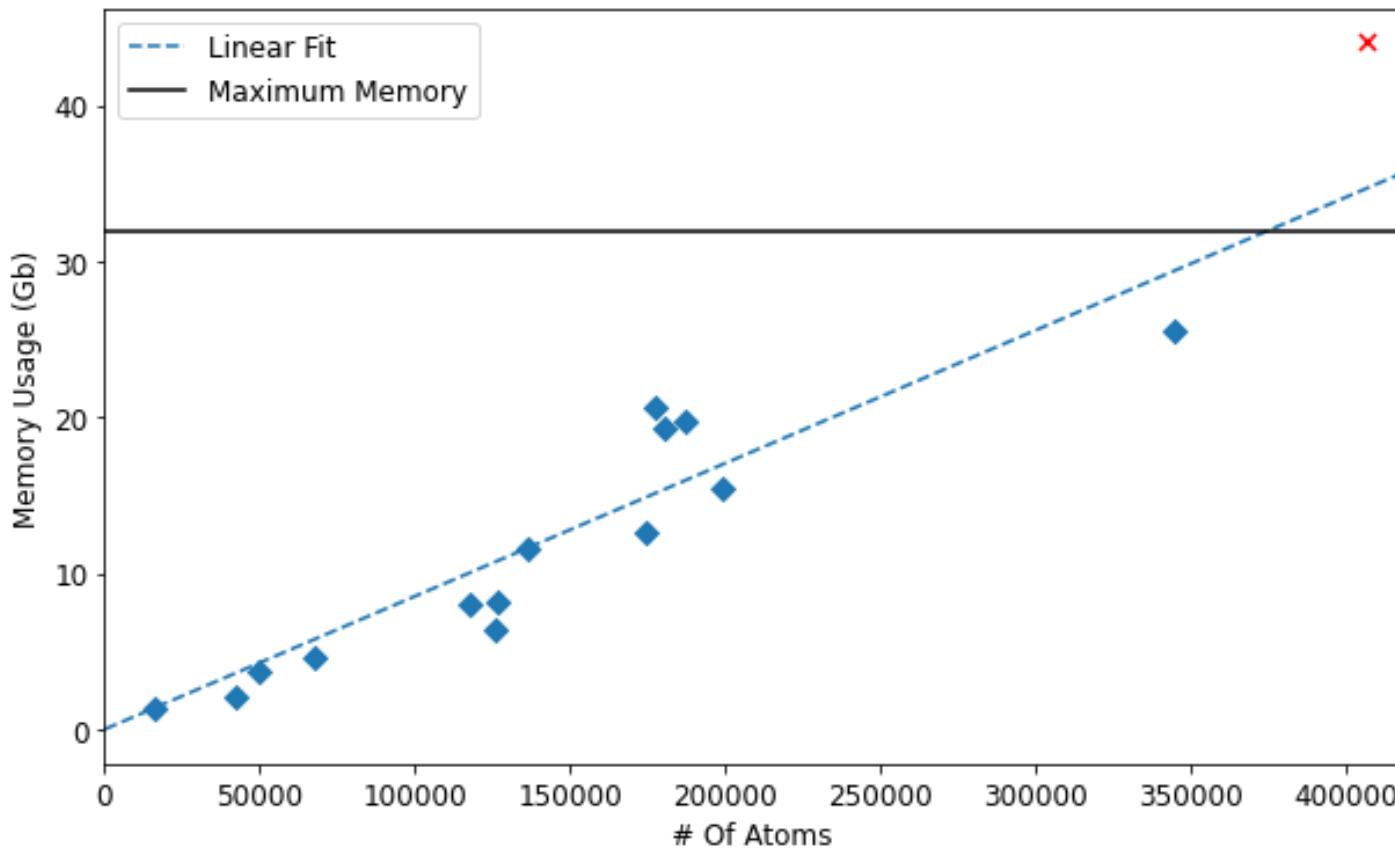
Step 3: Classification - Pyrococcus F. Encapsulin



Brown, Luque
In Preparation

Memory Limitations

Memory Required For Normal Mode Analysis



Brown, Luque
In Preparation

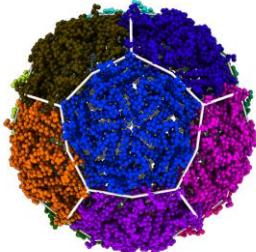
Limit of 32 Gb \approx 380,000 Atoms



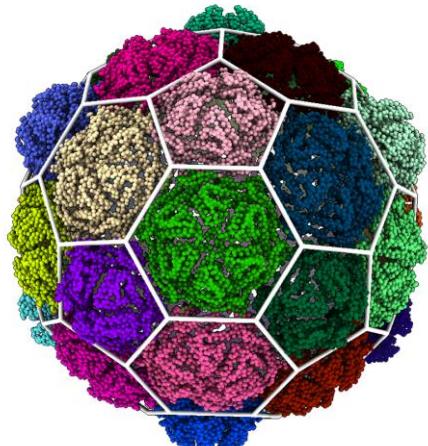
Isometric Phage T4
Exceeded Memory Requirements

Results – Hexagonal Structures

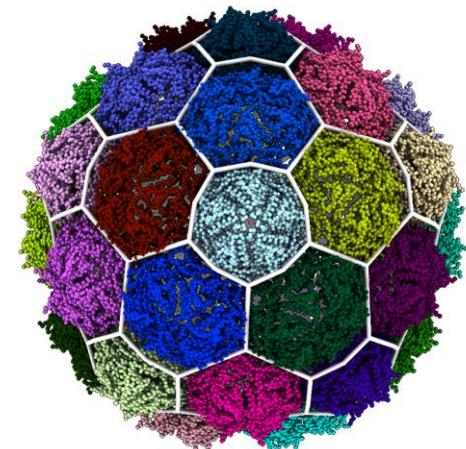
60 Proteins



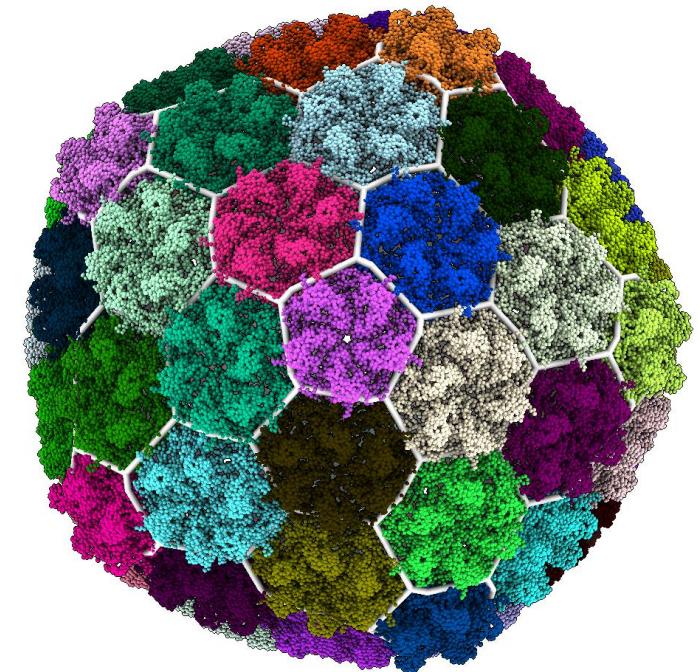
180 Proteins



240 Proteins



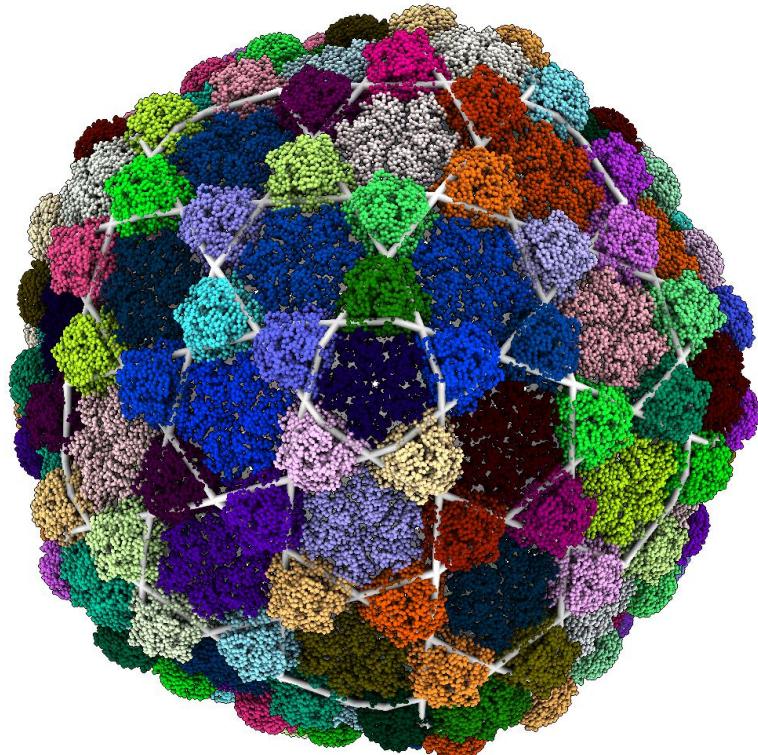
420 Proteins



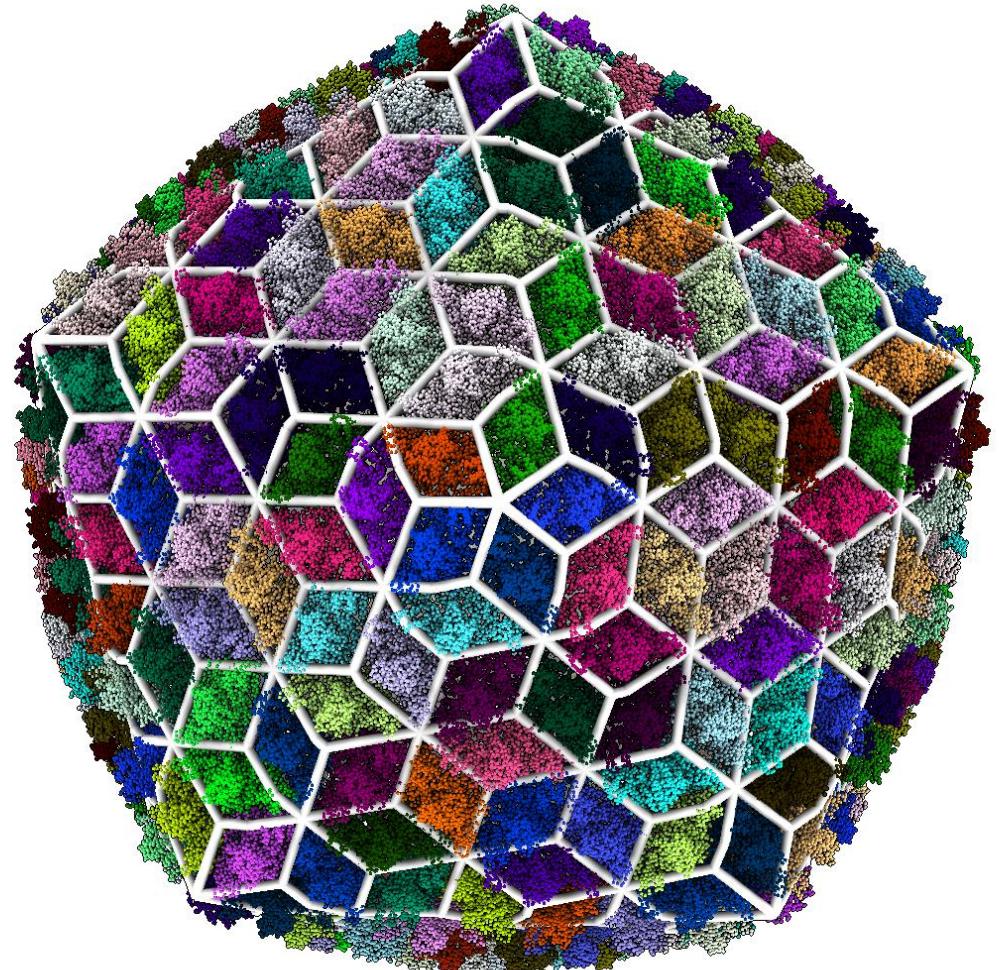
Encapsulins

Results – Reinforcement Proteins

420 Major Proteins
140 Reinforcement Proteins

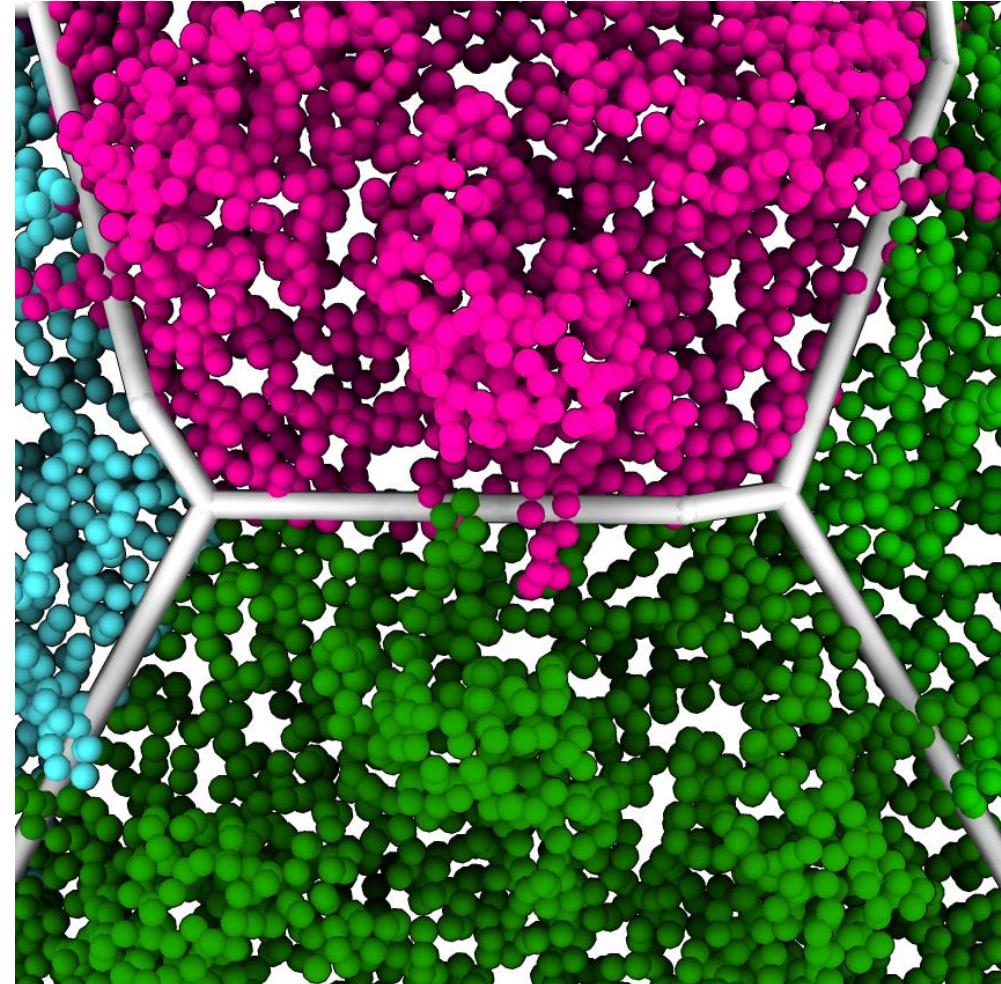
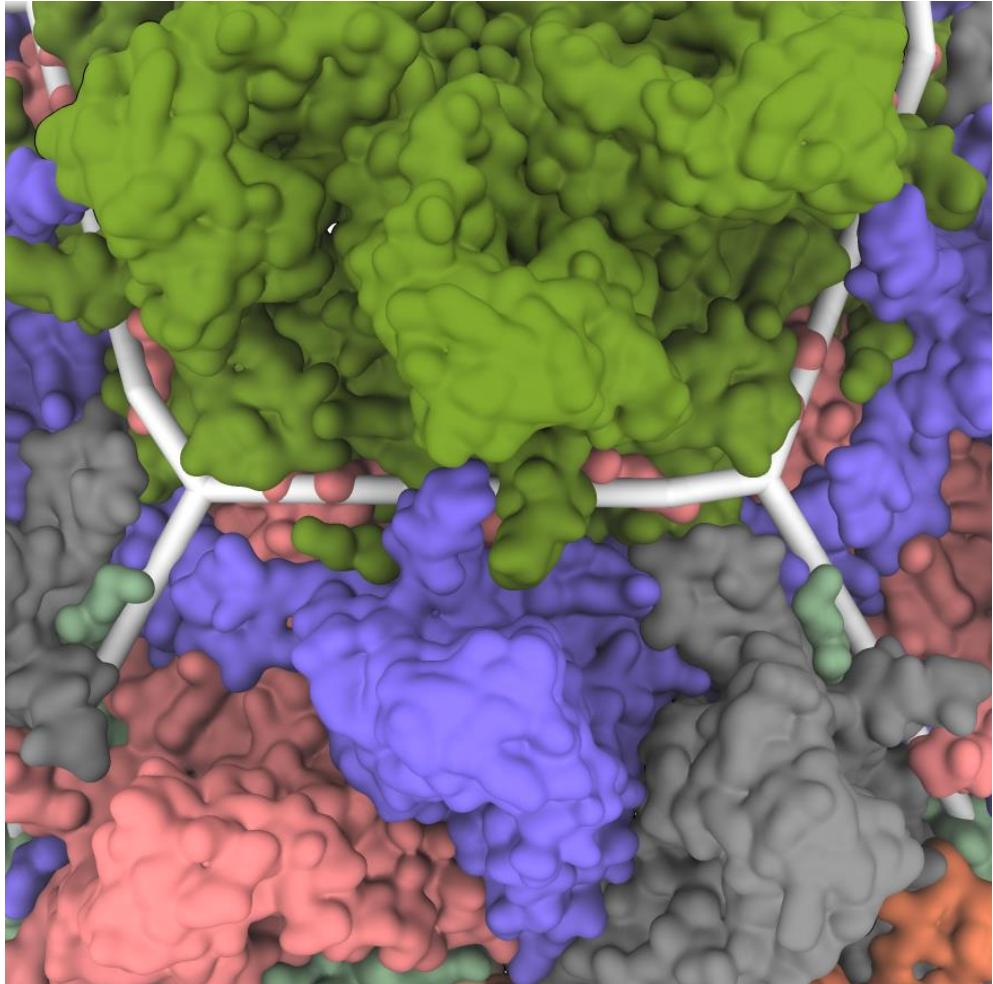


780 Major Proteins
260 Reinforcement Proteins



Results - Crossing Domains

- Bacteriophage P22



Conclusion

- HK97-Fold
 - Hexagonal Structure
 - Reinforcement Proteins
 - Extended Domains
- Classification Method
 - Larger Capsids - Reduce Memory Requirements
 - Symmetry
 - Modularity
 - Physical Models
 - Clustering Algorithms

Acknowledgements

Luque Lab - Capsid Classification

- Diana Lee
- Brandon Ricafrente
- Caitlin Bartels



PI

- Antoni Luque
Santolaria



UC San Francisco

- ChimeraX Development Team
- Tom Goddard



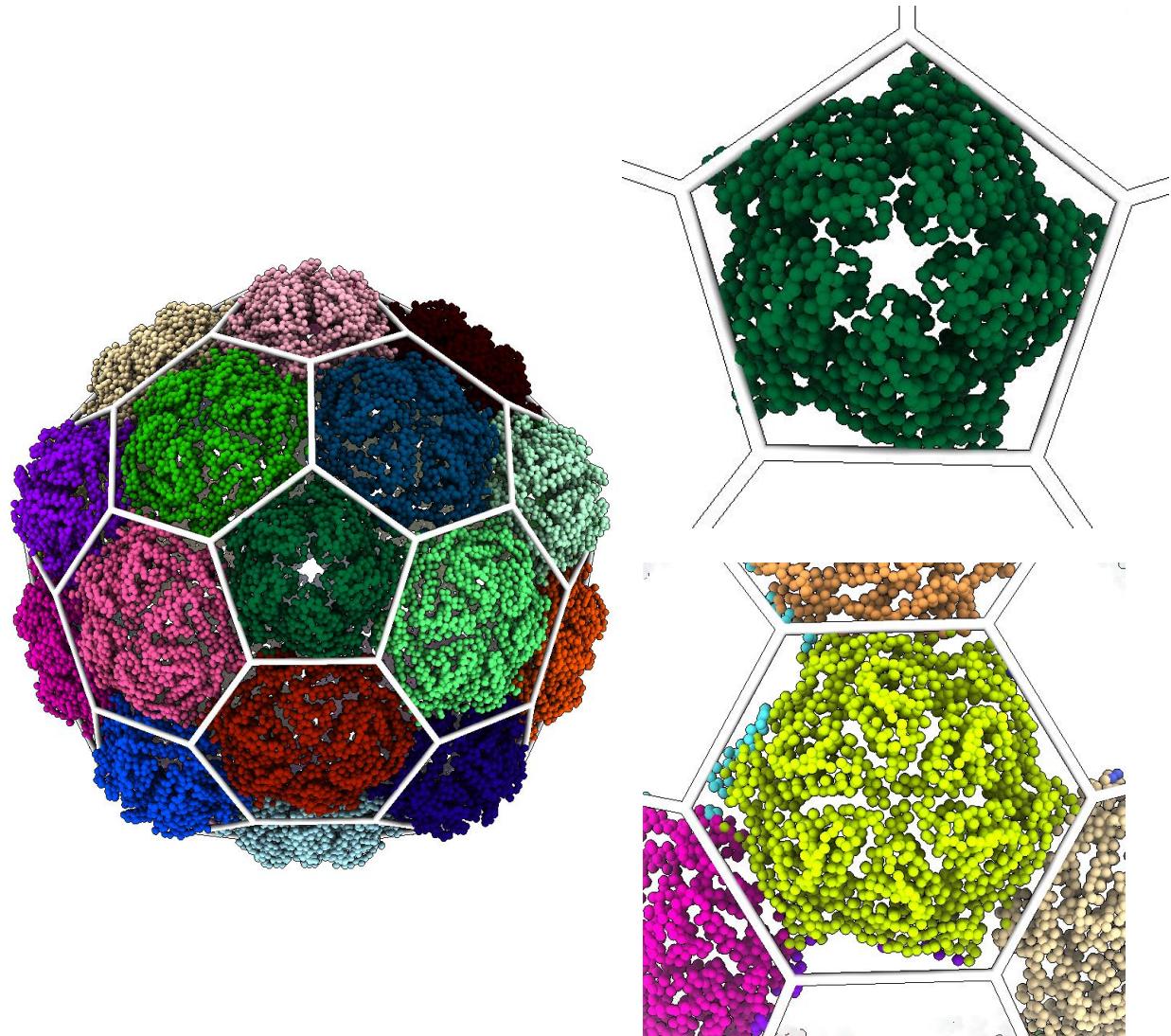
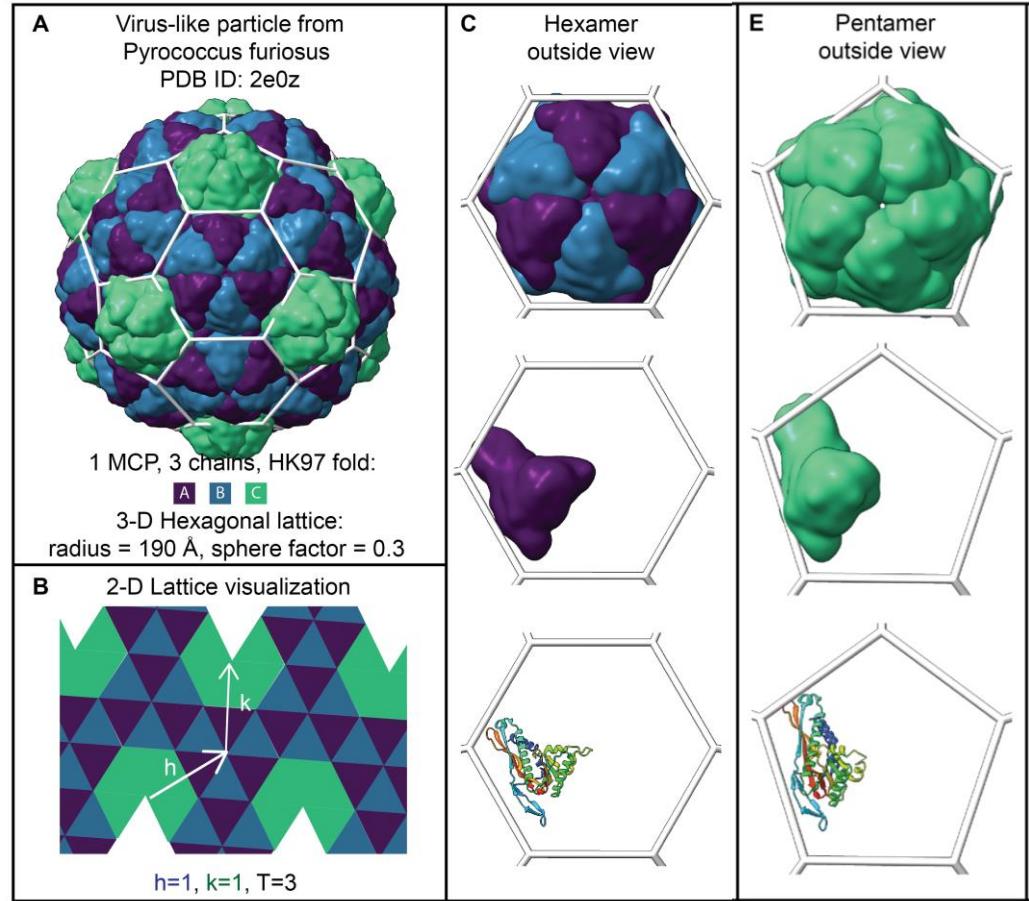
NSF Grant

- PI & Grant #



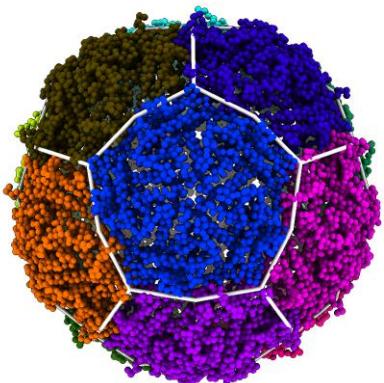
Questions

Results – Comparison With Geometric

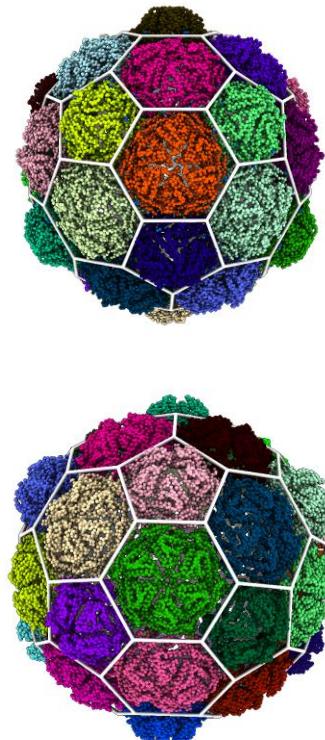


Results - Hexagonal

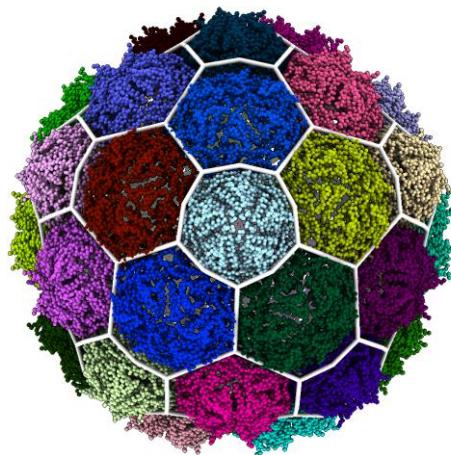
60 MCPs



180 MCPs



240 MCPs

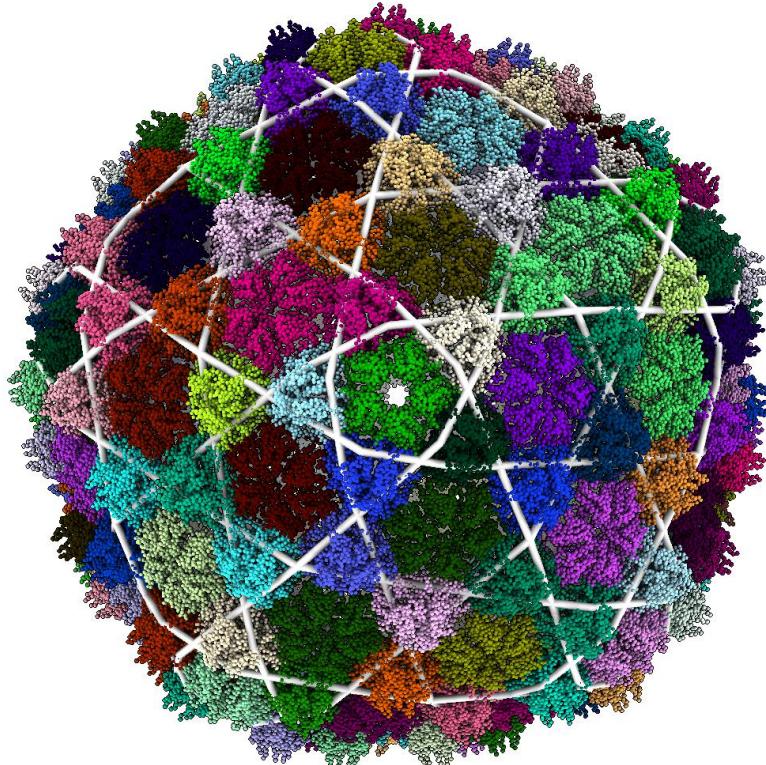
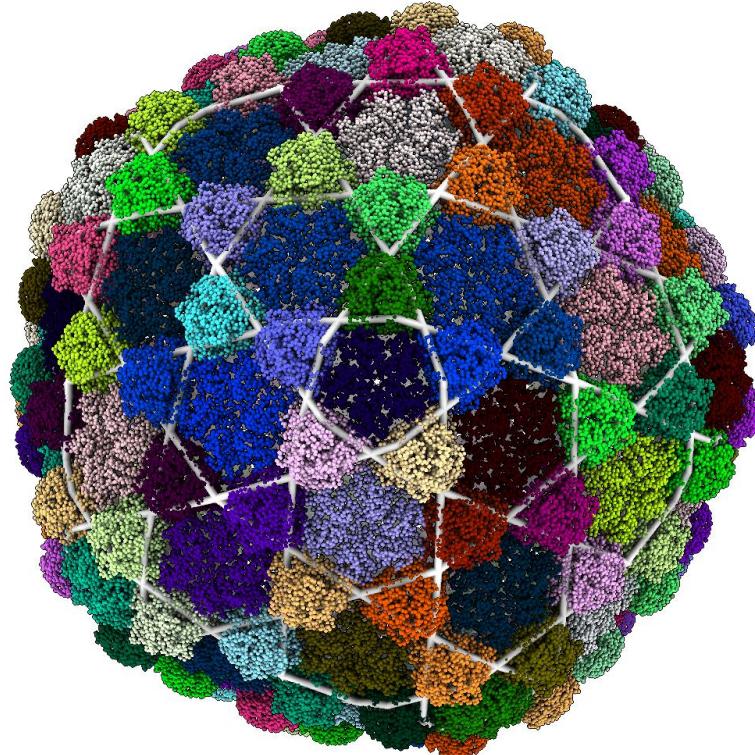


420 MCPs



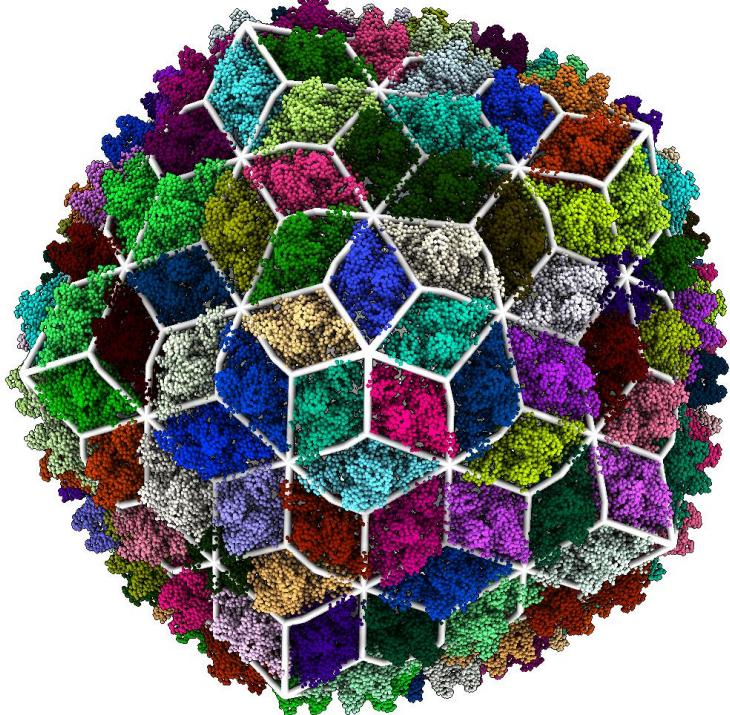
Results – Trihexagonal

420 MCPs & 140 Reinforcement Proteins

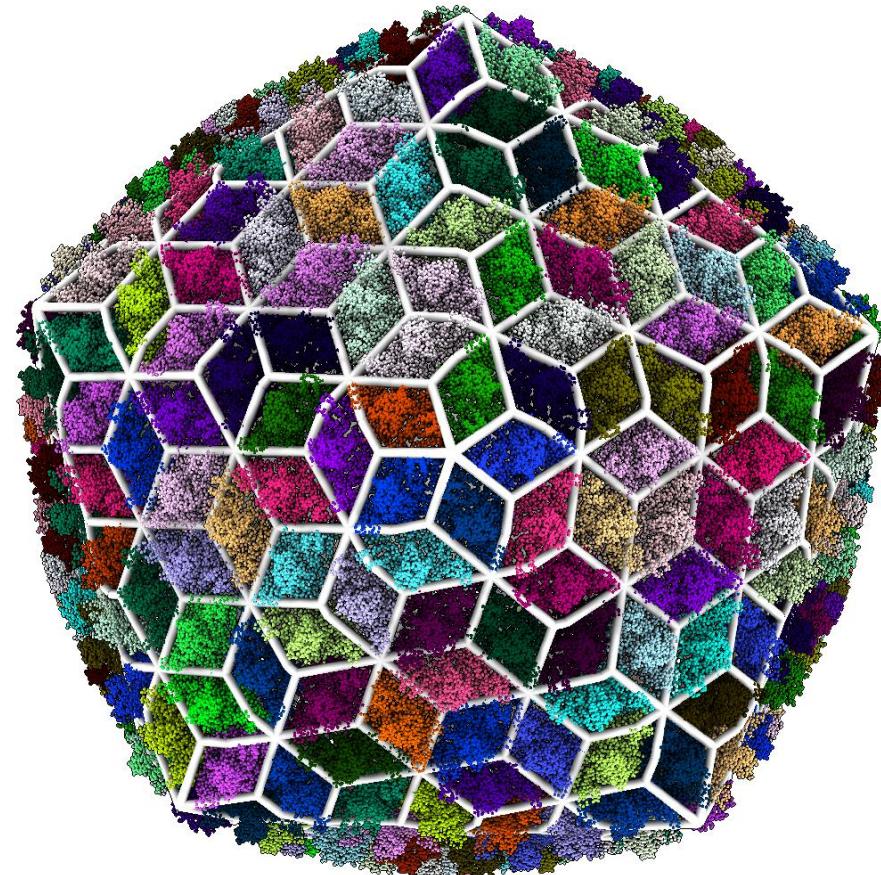


Results – Trihexagonal-Dual

420 MCPs & 140 Reinforcement Proteins

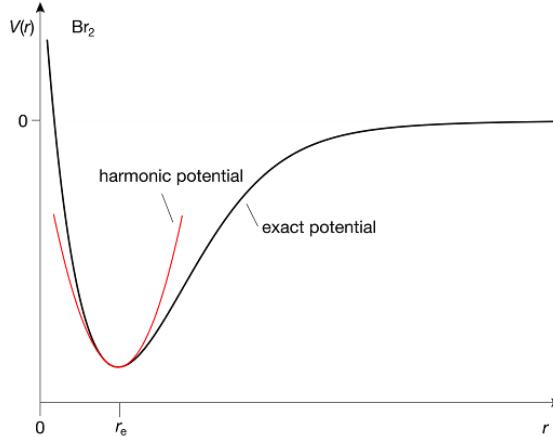


780 MCPs & 260 Reinforcement Proteins

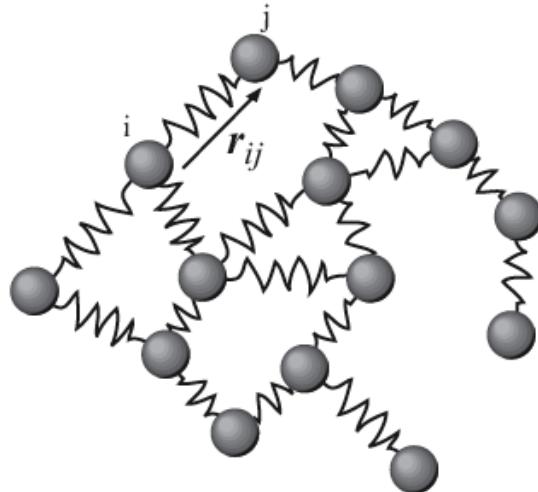


Model Assumptions & Limitations

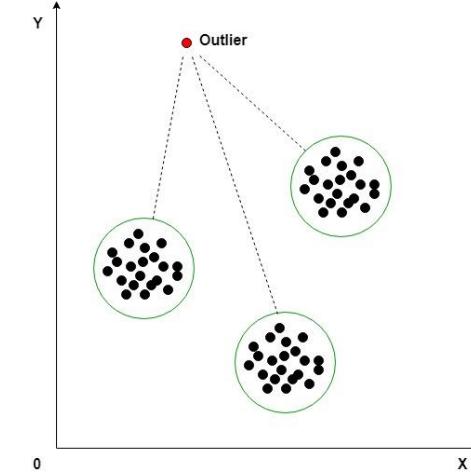
Harmonic Potentials



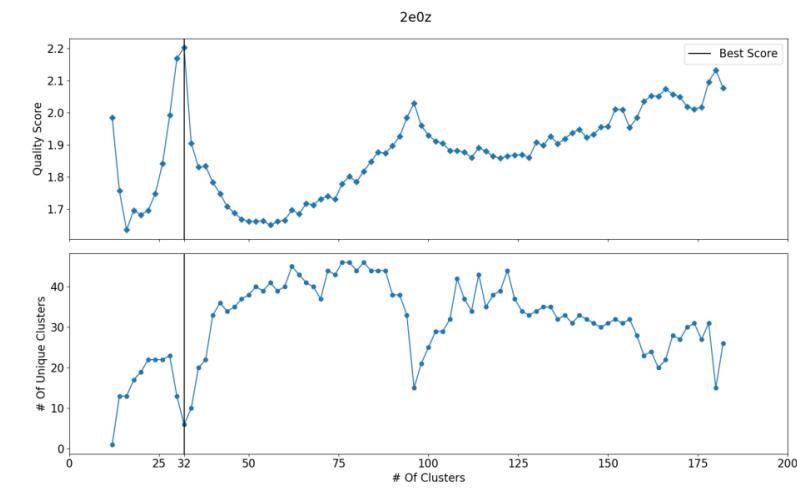
ENM
Coarse
Graining



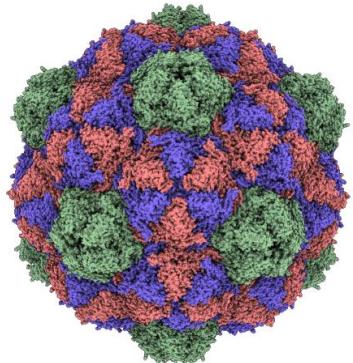
No Outlier Removal



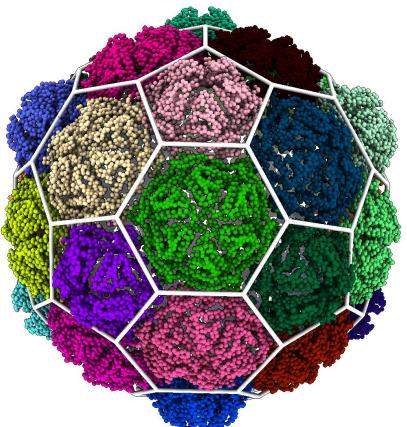
Subdivision
Selection



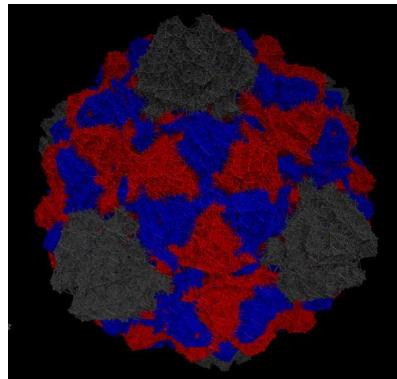
Atomic model



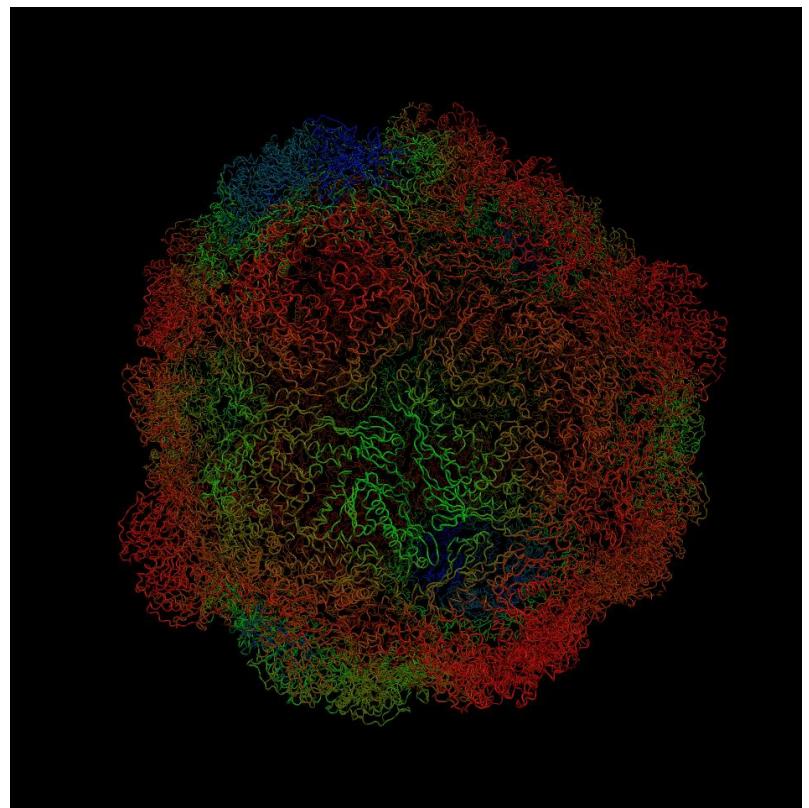
Icosahedral lattice assignment



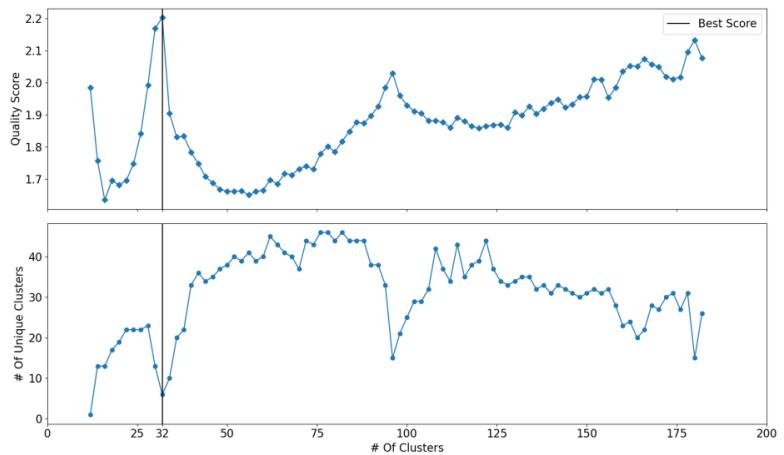
Elastic Network Model



Normal modes analysis



Scoring Subdivision



Performance

Constructing Hessian

$$\mathbf{H}_{ij} = -\frac{\gamma_{ij}}{(R_{ij}^0)^2} \begin{bmatrix} (x_{ij}^0)^2 & x_{ij}^0 y_{ij}^0 & x_{ij}^0 z_{ij}^0 \\ x_{ij}^0 y_{ij}^0 & (y_{ij}^0)^2 & y_{ij}^0 z_{ij}^0 \\ x_{ij}^0 z_{ij}^0 & y_{ij}^0 z_{ij}^0 & (z_{ij}^0)^2 \end{bmatrix}$$

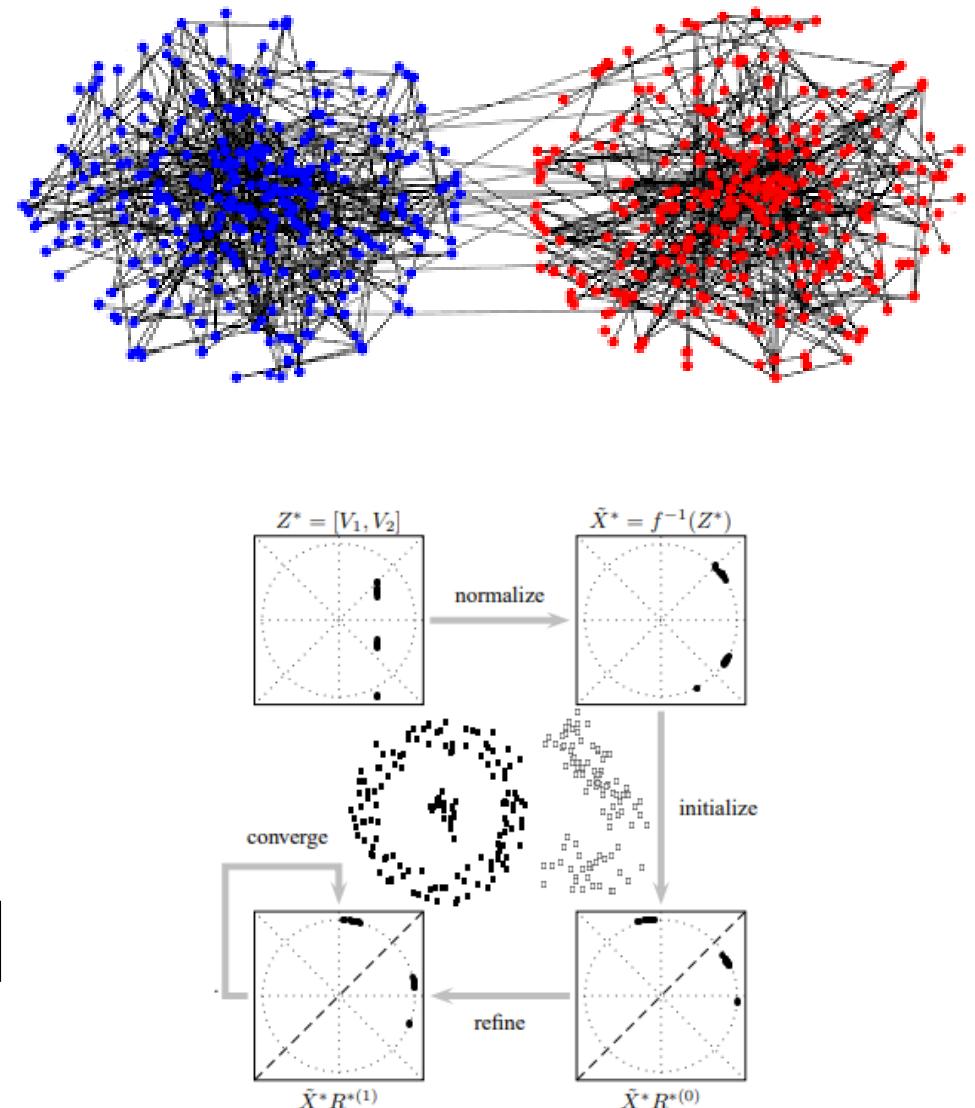
Calculate Normal Modes

$$\bar{H}^{-1} = \sum_{k=1}^{n_k} \frac{1}{\omega_k^2} v_k \otimes v_k$$

Of Modes

Eigenvectors/Normal Modes

Lowest Frequencies



Yu, Shi, IEEE 2003

Capsid	T #	Additional Proteins	Result
Th. Maritima	1	None	Hexagonal
Pyr. Furiosus	3	None	Hexagonal
Myx. Xanthus	3	None	Hexagonal
Quasi. Thermo.	4	None	Hexagonal
P22 Mature	7	None	Hexagonal
Phage Sf6	7	None	Hexagonal
Phage T7	7	None	Hexagonal
Bordetella	7	2-Fold Axis	Trihex-dual
Siphov. Mic1	7	2-Fold Axis	Trihex-dual
Siphov. TW1	7	3-Fold Axis	Trihexagonal
YSD1 Phage	7	3-Fold Axis	Trihexagonal
Siphov TW1	13	2-Fold Axis	Trihex-dual

Quasi-Rigid Domain Decomposition

- We implement the method described in (Ponzoni & Polles 2015) to identify quasi-rigid domains
 1. Represent the capsid using an Elastic Network Model (ENM)
 2. Apply Normal Mode Analysis to approximate the pairwise distance fluctuations between residues
 3. Use Spectral Clustering techniques to optimally divide the capsid into quasi-rigid domains
 4. Visualize the result and assign a geometric classification based on the clusters